DOCUMENT RESUME

ED 107 716 TM 004 563

AUTHOR Korotkin, Arthur L.
TITLE The Evaluation of Dropout Prevention Programs.
INSTITUTION American Institutes for Research in the Behavioral Sciences, Washington, D.C.
REPORT NO 44900/FR-1
PUB DATE Dec 74
NOTE 50p.

EDRS PRICE MF-$0.76 HC-$1.95 PLUS POSTAGE
DESCRIPTORS Disadvantaged Youth; Dropout Programs; *Dropouts; Educational Innovation; Educational Objectives; Elementary Secondary Education; Evaluation Methods; Evaluation Needs; Feedback; *Formative Evaluation; *Literature Reviews; Measurement Techniques; Occupational Guidance; Program Effectiveness; *Program Evaluation; *Summative Evaluation

ABSTRACT

The meaning of program evaluation in the context of educational innovation has been expanded to embrace all aspects of a program, both formative and summative. This review seeks to identify meaningful literature on programs related to dropout prevention and the preparation of the disadvantaged for the world of work through vocational training, educational enrichment, and counseling. Only 20 of 100 evaluations of related programs were deemed sufficiently objective for inclusion in this review. Measures used to evaluate related programs ranged from IQ, aptitude, and achievement to personality, attitudes, and interests. It was concluded that there is little in the way of evaluation approaches, techniques, or measures that is not currently incorporated or accounted for in the existing evaluation system designed by Gibboney and Associates.
(Author/BJG)

ED107716

44900/FR-1

THE EVALUATION OF DROPOUT PREVENTION PROGRAMS

SCOPE OF INTEREST NOTICE
The ERIC Facility has assigned this document for processing to: TM CG
In our judgement, this document is also of interest to the clearinghouses noted to the right. Indexing should reflect their special points of view.

By

Arthur L. Korotkin, Ph.D.

American Institutes for Research
3301 New Mexico Avenue, N.W.
Washington, D.C. 20016

December 1974

TM 004 563

ACKNOWLEDGEMENT

The author gratefully acknowledges the contributions of Dr. Paul A. Schwarz, President of the American Institutes for Research and Director, Social Change Division, and Dr. Robert E. Krug, Deputy Director, Social Change Division, to the preparation of this report. Their experience and insight with regard to the evaluation process was invaluable in attempting to structure what has already been done and, more importantly, to examine what can be done.

A. L. KOROTKIN

TABLE OF CONTENTS

I	Introduction	1
II	Problem	3
III	OIC/CIP Program Evaluation	5
IV	The Measures	8
	A. Previous Comprehensive Evaluation Effects	9
	B. Results for Previously Used Measures	23
	1. Measures of Progress and Immediate Outcomes	23
	a. Continuing Education - Retention in School	23
	b. Continuing Education - Attendance	24
	c. Intellectual Achievement	24
	d. Basic Academic Skills - Achievement Tests	24
	e. Basic Academic Skills - Grades	26
	f. Measures of Attitude and Alienation	27
	g. Acquisition of Social Skills	29
	2. Post Program Objectives and Long Range Goals	29
	a. Continued Education and Employment	29
	b. Citzenship and Community Concern	31
	c. Other Long Range Effects	31
V	Summary and Conclusions	32
VI	Appendix A	37
VII	References	43

I. INTRODUCTION

The approach to evaluation has been modified and refined over the past several years. Even the basic concept has changed from one of assessing "program success" to that of a process which is totally integrated into the program itself and attempts to contribute to "program success". It is no longer sufficient for the evaluator to arrive at the end of a project and using his traditional measurement tools and statistical techniques assess the success or failure of the program. This unitary, after-the-fact approach has in most cases produced quite unsatisfactory results which were useless to both those who needed to make management decisions about the program and to those who were involved in the program development. In his opening remarks to an invited seminar on evaluation, the late Dr. S. Rains Wallace expressed his concern that meaningless evaluation was in fact ruining the cutting edge of innovation. He stated also that, "Perhaps most importantly, many of us share an uneasy but firm conviction that our evaluation problems cannot be solved with the traditional scientific methods and paradigms and the statistical refinements we have developed for their analysis and interpretation" (American Institutes for Research, 1970, p. VI).

It would appear that the current emphasis on innovation and social change has prompted a new look at evaluation. There is an increasing need to be able to demonstrate that programs or parts of programs are indeed effective or at least that they can be made effective. There is a need for those implementing a program to be able to derive feedback from what they are doing so that changes can be made in a timely and effective manner before the program ends. There is certainly a need for those in management positions to be able to know if progress is being made or if a program has achieved the desired results. Thus the technology of innovation has brought with it a parallel development in the technology of evaluation. An attempt is being made to deal with the problems of assessment in a sophisticated but practical manner so that the results are both valid and useful.

A new vocabulary has developed. Terms such as objectives, outcomes, long-range goals are used to specify the desired results of a project, while terms like "process evaluation", "product evaluation", "formative evaluation" and "summative evaluation" are used to describe the targeted areas for the various components of the overall evaluation program. The concept of linking a measure or instrument to a meaningful reference point, group or situation (criterion-reference) has received recent support for assessing change. There is continuing research on how best to examine the real ultimate impact of a new program or innovation. This entire evaluation technology effort is embedded in the problems of working in the "real world". Reality brings with it the problems of limitations in dollars and time as well as various practical, social and political barriers to implementing the desired program and/or evaluation plans.

Our increasing level of knowledge and awareness in both innovation and evaluation has resulted in a more comprehensive integrated approach to both. In addition, the identification of objectives not only aids the evaluation effort but it also requires that those who are implementing the program define their goals more precisely. The data from the Formative and Process Evaluations can be fed back to those responsible for the program so that changes can be made in the methods or resources applied. Therefore, both the innovation and evaluation programs are tied together to their mutual benefit. Evaluation has been broadened to include an assessment of the ongoing processes and components of the program, and sharpened to include the measurement the degree of success in attaining certain precise objectives and outcomes.

II. PROBLEM

The goal of this project was to enhance the existing evaluation of the OIC Career Intern Program by Gibboney and Associates by the identification and recommendation of supplementary criterion-referenced measures. These recommended measures would be aimed at assessing both immediate and long-term program goals. The procedure for identifying these measures were:

A. A review of existing literature in the area of dropout prevention, career intern programs, and other related areas.
B. An analysis of the objectives and goals of the OIC Career Intern Program.

It was intended that the recommended set of measures be consistent with the established evaluation and objectives, be supplemental rather than redundant to measures already included in the evaluation program, and be sensitive to CIP program outcomes.

The results of the literature review are presented in a separate report by Dr. David V. Tiedeman and Dr. Anna Miller-Tiedeman entitled "Career Initiation When Alienation from Secondary Schools Occurs: A review of the literature." This survey of the literature contains over 100 references to related programs in dropout prevention and the preparation of the disadvantaged for the world of work through vocational training, educational enrichment, and counseling. It highlights the different approaches taken and discusses the results of some of the more significant projects.

Using the references identified in the literature review, each report was examined carefully to determine if any type of evaluation was conducted as part of the study. For those projects having an evaluation component, they were again reviewed to identify the specific objectives, and the measures used to assess the program outcomes. It should be noted that the number of programs employing identifiable measures were considerably less than the 100 references in the literature review.

Because of the limited data available from related programs the measures described in this document are not limited to criterion-referenced techniques as originally intended but include the entire spectrum of measures used in the various programs. It was hoped that such a broadening of the scope would increase the utility of this document. For the most part the measures tend to be the obvious ones and are already accounted for or included in the Gibboney evaluation package. There were some unique measures identified and these will be discussed in the following sections within the framework of a generalized evaluation model.

The literature itself, as typical of most of the social science literature, tended to be quite variable with regard to both quality of the research and quality of the reporting. Only a limited number of the projects reported in the literature attempted any type of objective evaluation. Very few of these studies precisely described their assessment effort; and fewer described, to any useful level of detail, the characteristics of their objectives, measures, or subject populations. In the small number of cases where the descriptions were adequate, the subject populations tended to be so limited in size and scope, and the measures so program specific, that the results could at best be described as inconclusive and any attempt at generalization made with only the greatest of caution. Therefore, despite the over 100 references to related programs in the literature review, objective data on various kinds of measures could only be abstracted from the 20 studies cited in this report.

As a final note regarding the literature in this area it should be recognized that most of the studies of interest were reported in the so called "fugitive" literature. This means that the information about related programs and studies were reported only in various project reports as opposed to formal archival journal literature. The location and acquisition of such documents was difficult and time consuming at best and in some cases proved impossible.

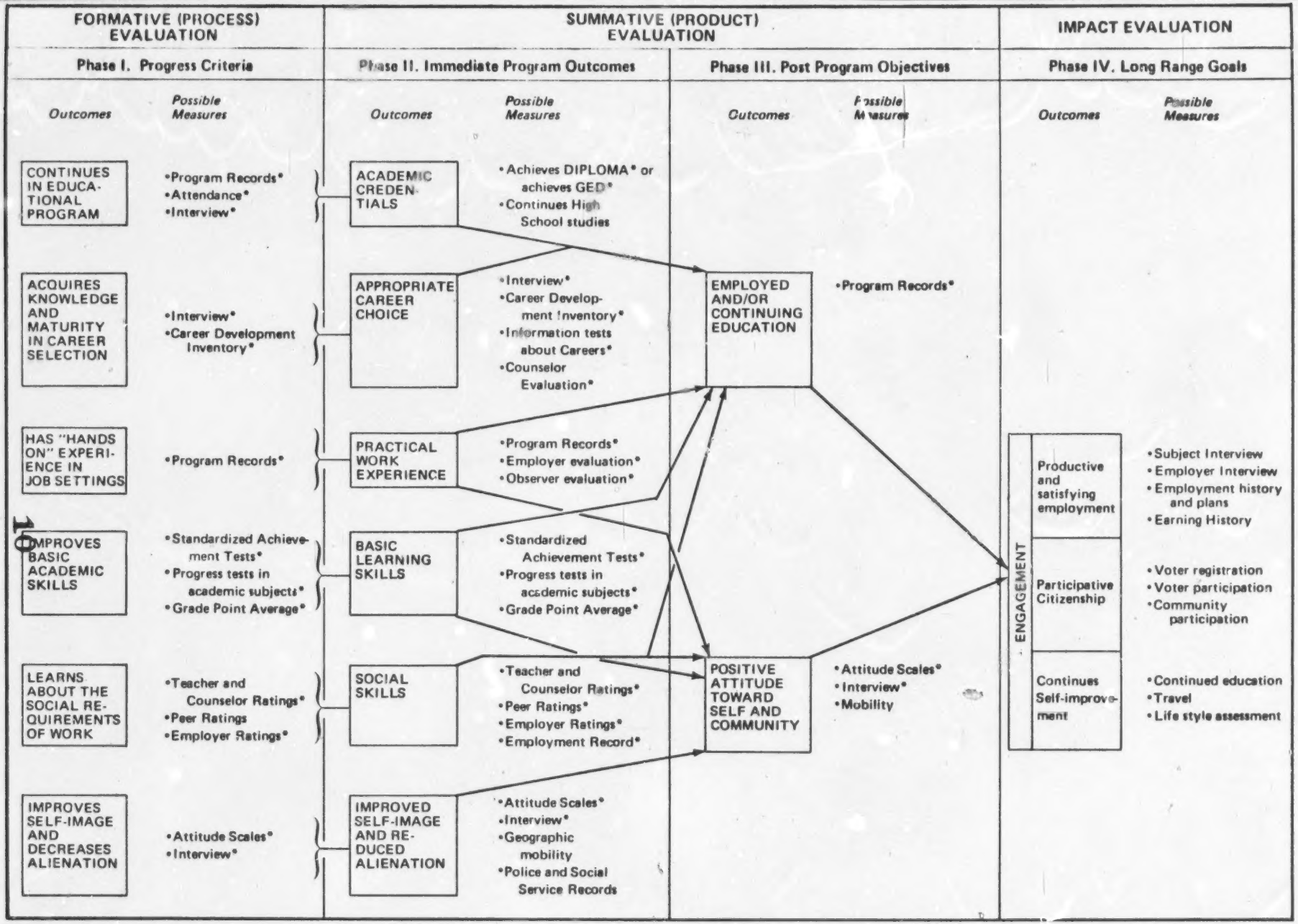
III. OIC/CIP PROGRAM EVALUATION

In order to place the data from the literature in perspective and to determine the requirements for additional assessment measures, it was necessary to construct a generalized model of the OIC/CIP Program evaluation. The model presented here is not necessarily congruent with the one that underlies the current evaluation effort. It was developed only to serve as a frame of reference for the following discussion.

After site visits, discussions with program staff and the review of a considerable amount of program materials, the following six objectives were identified as the desired major program outcomes.
To assist the students in:

1. obtaining academic credentials (Diploma or GED)
2. making an appropriate career choice
3. having a variety of practical "hands-on" work experience
4. developing a set of relevant basic academic skills
5. acquiring a set of social skills required for job success
6. improving their self image and reducing alienation

The evaluation effort is viewed as having four (4) phases. The interaction of the program outcomes and the various evaluation phases are represented in Figure 1. Possible measures are also shown for each outcome. Phase I is a Formative or Process Evaluation phase for assessing program progress. This is conducted while the program is ongoing and is aimed at measuring the progress being made in the various components. Such information can be fed back into the program to effect improvements. The second and third phases are Summative or Product Evaluations. Phase II is directed at "Immediate Program Outcomes" while Phase III examines the "Post Program Objectives" or longer range program effects. The Phase II outcomes are the end points of the Phase I progress evaluation, i.e. it is keyed to the same general outcomes. The object of Phase II is the evaluation of the effectiveness of the program or the program elements, while Phase III looks at the broader questions of program scope--is the program doing the right things? Is it doing enough?



* *Currently in use or available.*

Figure 1. A GENERALIZED PARADIGM FOR THE OIC/CIP PROGRAM EVALUATION

It is expected that by the Phase II evaluation the program participants have:

1. completed the program and have received their high school diploma or GED.
2. made an appropriate choice of careers which is mature, realistic and achieveable, and based upon useful knowledge about that career gained in the program.
3. participated in a variety of practical experiences in various settings.
4. attained an acceptable level of achievement in such basic academic skills as reading and math.
5. recognized and developed the necessary social skills required for successful employment.
6. developed a good self-image, self-confidence and reduced feelings of alienation from society and its institutions.

All of these outcomes feed into those for the Phase III evaluation. Phase III is conducted some time after the end of the program in a month, six months or a year. Obviously it can also be conducted several times at various time intervals after the end of the program. It addresses the two major expected outcomes:

1. Is the program participant employed, continuing his education or training, or a combination of the two?
2. Does the program participant have a positive attitude toward himself and his community?

The measures here can be some of the same ones used in Phase I and II; however, some new ones may have to be introduced. It may be necessary to examine surrogate indicators of the desired outcomes. For example, geographic mobility may be an indicator of self-confidence and decreased alienation. Promotions on the job, higher earnings than equivalent program non-participants, participation in community activities may all be indicators of the [illegible]ed program outcomes.

Finally, Phase IV looks at the Long Range Goals of the program and attempts to assess its ultimate impact. The question addressed at this point (some time in the future) is not whether program objectives were achieved, but whether or not they were the right objectives. The ultimate criterion for program success must be to determine if the individual participant is better off for having been in the program. This, obviously, cannot be done until long after the program ends, although some interim progress measures are certainly possible and desirable.

Again such surrogate measures as voter registration, community participation, continued education and assessment of life style may be indicative of the exacted long range goals of:

1. Productive and satisfying employment.
2. Participative citzenship.
3. Continued self-improvement.

These three goals all contribute to the individual's "engagement" with society (Seeman, 1972).* They are difficult to measure because they are removed from the basic program outcomes in terms of both time and level of abstractness. However, one possible approach to measurement is to assess engagement by looking at the investment the individual is willing to make in himself and/or his community. Such investment may be in terms of time, effort and money (Krug and Jong, 1974). The investment may be a way of quantifying comitment and thus "engagement". Unfortunately there is little in the literature which reflects on this phase of evaluation. If measures are to be used in Phase IV, then an effort will have to be undertaken to develop and validate them.

*The term "engagement" is used here by Seeman as an antonym to alienation.

IV. THE MEASURES

The measures used to evaluate the related programs identified in the literature ran the gamut of conceivable assessment tools. They included measures of IQ, aptitude, achievement, personality, attitudes, and interests. In addition, such program related measures as attendance measures, grade point average, rating scales and interview forms were used. A complete list of those measures identified in the literature is presented in Appendix A by category and article referencing the particular measure. Virtually all of the measures fall into the Phase I and Phase II evaluation categories, i.e., they are measures of progress and immediate program outcomes. Most of them were given at the beginning, during and at the end of the various programs. Very few measures for Pahse III (Post Program Objectives) and Phase IV (Long Range Goals) were identified.

Only a small number (11 to be exact) of the related studies identified included any type of comprehensive evaluation effort. Those that did take a broader approach to evaluation tended to be similar to each other in that they attempted to assess the same general areas of progress and immediate program outcomes. The program objectives most frequently examined were the assessment of:

1. Program effectiveness on overcoming deficits in academic skills, particularly reading and arithmetic.
2. Personality and attitude factors associated with improving the participant's self-image.
3. Changes in feelings of alienation from the main stream of society and its institutions; e.g., school and the world of work.
4. The acquisition of certain social skills necessary for holding a job.
5. The acquisition of technical skills needed for obtaining and holding a job.

Similarity between the goals and objectives of the OIC/CIP program is obvious. It should be noted, however, that there is a major difference between the OIC/CIP program and other programs identified. This difference lies in the combined career education/work experience approach taken in the CIP program. While other programs did include personal and career counseling

as part of their multi-faceted approach to the dropout problem, we were unable to identify another project which had a dual emphasis of work experience with career education and guidance. Thus, there was little in the literature regarding the assessment of the maturity and level of knowledge with which career decisions are made.

A. PREVIOUS COMPREHENSIVE EVALUATION EFFORTS

Described below are a series of studies which approach the evaluation problem in a comprehensive manner. These evaluation efforts are described and the measures identified. In most cases the measures were used as pre and post program tests. The actual results associated with these measures, if any were reported, will be found in the following section on "Results from Previously Used Measures".

Hamburger (1965) reporting on a program designed to aid disadvantaged youth with a co-op work study program reported on data collected from 108 Co-op students and 91 Controls. The data were collected in 1962-1963. His measures included:

1. Stanford Achievement Tests, Advanced Battery, Form JM, Arithmetic Reasoning and Reading Subtests.
2. Otis Intelligence Tests, Self-Administered, Gamma Form FM. This test was selected because of its similarity to the typical Civil Service examination.
3. The Differential Aptitude Test Battery, Form A, Clerical Test. This test was selected because it was especially relevant to the large number of cooperative students who entered the office work area.
4. The Work Inventories developed by Dr. Donald E. Super and Associates. This measure was used to assess changes in work attitudes and values.
5. The Life-Planning Questionnaire developed by Morton Hamburger. This measure was used to compare and contrast levels of aspiration and levels of expectation in the vocational and educational areas.
6. The Incomplete Sentence Blank developed by J. V. Rotter, High School form. This instrument was used as a personality screening device.

7. Job ratings were obtained from each of the cooperative students.
8. Interview data using a structured interview schedule were obtained on each of the students.
9. The Revised Occupational Scale for measuring socio-economic status, developed by Hamburger, was used to rate levels of occupational aspiration and expectations.
10. Additional data on the students were obtained from school records. These included absences, lateness, discipline slips, grades, and school tests.

Leubling and Trobe (1965) reported on a program which focused on the vocational, educational, and personal problems of school dropouts between the ages of 16 and 18. Their evaluation program consisted of the following:

1. Wechsler Intelligence Scale
2. Monroe Silent Reading Test
3. Woody-McCall Arithmetic Fundamentals
4. Minnesota Clerical Test
5. Vocational Advisory Service Simple Posting Test (Program Specific)
6. Minnesota Spatial Relations Test
7. O'Connor Wiggley Block
8. O'Connor Finger and Tweezer Dexterity
9. Minnesota Rate of Manipulation (placing and turning test)
10. Vocational Advisory Service "Information Schedule" (a projective questionnaire developed for the project)

The following tests were also used on a selective basis:

1. The Thermatic Perception Test
2. Bender-Gestalt Test
3. Draw-a-person Test
4. Porteus Mazes
5. Kohs Block Design
6. Stenquist Mechanical Assembly
7. IER Girl's Mechanical Assembly Test
8. Pennsylvania Bi-Manual Test
9. Vocational Advisory Spelling Test
10. Gates Reading Test

11. Peabody Picture Vocabulary
12. Otis Arithmetic Reasoning

Of the 263 clients, 218 were administered a full battery (158 were male and 60 were female). However the test battery was modified when necessary to consist only of non-verbal tests since many of the clients had language and reading deficiencies. In addition, many of the clients had negative feelings about anything resembling an academic pencil and paper test so that only performance tests were used under those circumstances as well.

In "A Pilot Project to Develop a Program of Occupational Training for School Alienated Youth" (1967), a comprehensive evaluation program was attempted. It was, however, limited to the assesssment of progress and immediate outcomes. The experimental design included the use of a control group so that unwanted extraneous effects that result from normal growth and the effects of factors other than the program could be considered in the analysis. However, it was impossible to obtain a random sample of control students and the sample size was limited to fifteen. Therefore, comparisons between the experimental and control groups had to be made with extreme caution. Data analysis of the experimental group itself showed N's ranging from 7 to 39.

The measures used in the evaluation are listed below:

	Measure	Characteristics or Variables
* 1.	California Study Method Survey	School adjustment, study habits, skills
* 2.	California Test of Personality	Personal and social adjustment
3.	IPAT Anxiety Scale	Anxiety level
* 4.	Lorge-Thorndike Intelligence	Academic aptitude
5.	Brown-Carlsen Listening Comprehension Test	Ability to comprehend ideas and to remember significant details of listening
6.	STEP Science Test	Understanding general science
7.	STEP Math Test	Understanding general mathematics
8.	STEP Social Studies Test	Understanding social studies
* 9.	California Reading Test	Reading vocabulary and comprehension
10.	Minnesota Vocational Interest Inventory	Occupational interest

*11.	NYU Speaking Test	Clarity in reading and conversation
*12.	Occupational Reading Scale	Work attitude in occupational skills
*13.	Case Study	Self-concept, social behavior, attitude toward society, school adjustment, and work attitude

It was planned that except for the case study and the occupation rating scale, all instruments were to have been administered prior to entering the program and upon leaving. However, the original test battery proved to be too long and cumbersome and the battery was reduced from 13 items to the seven instruments marked with an asterisk in the above list.

Kaufman, et. al., (1968a) conducted a study to examine the relative effects of skill training versus a diploma program for a group of high school dropouts. There were 60 students in the Diploma Group, 28 in the Skill Training Group, and the Control measures were taken on from 20 to 63 students with actual group size depending on the test. The measures used to evaluate the effects of the program were:

	Measures	Characteristics or Variables
1.	OTIS Quick Scoring Mental Ability Test, GAMMA Test Forms EM And FM	A traditional verbal IQ test
2.	Revised BETA Examination	A non-verbal intelligence test consisting of various performance subtests
3.	Culture Fair Intelligence Tests (Forms A and B)	A non-verbal intelligence test using relationships between geometric figures
4.	Edwards Personal Perference Schedule	Measures of 15 personality variables
5.	The Gough Adjective Checklist	A personality measure by which an individual describes himself using any of 300 common adjectives
6.	Coopersmith Self-Esteem Inventory	A measure of self-esteem in which the respondent checks off whether

	the various statements described are like or unlike himself
7. Minnesota Vocational Interests Inventory	Comparison of respondents interests to those of current members of various occupations
8. Stanford Achievement Tests (Advanced paragraph meaning test Forms W and X; Advanced Arithmetic tests W and X)	Standardized acheivement tests in reading and arithmetic

Kaufman used the standardized measures to assess immediate program effects on intelligence, personality, vocational interests, and academic achievement in reading and arithmetic. However, no long term follow-up was made.

Erickson and Hamler (1972) reported on the results of the Benjamin Franklin's Urban League Street Academy which was set up to deal with school alienated youth. The major objective was to "... help dropouts to return to school or to enter the world of work prepared and motivated to adjust and achieve satisfactorily." In order to do this the Academy was prepared to provide help in:

- Overcoming academic and vocational weaknesses to enable the dropouts to continue their education.
- Overcoming alienation from self and others, from the educational settings, and from school personnel.
- Remediation of below level academic skills.
- Developing, and acting upon, self-confident attitudes particularly by participating in continued education.
- Developing positive attitudes toward self, school, and society.
- Developing a mutual trust, communication skills, and legitimate concern for their community.

The evaluation measures used were:

Measures	Characteristics or Variables
1. School records, interviews with school personnel, parents, and community	School and employment status
2. Stanford Achievement Test	Achievement in English and Math

3. Questionnaire Data	Measure of alienation
4. Actions to continue education	Development and application of self-confidence
5. Unstructured interview with community, students, parents and staff	Development of trust, communication, and concern for the academic program and the community
6. School records	Cost per student

A unique feature of this particular evaluation was the inclusion of measures of self-development. The reduction of alienation and development of self-confidence are items that were used in other evaluations. However, the attempt to assess the ability to act on that self-confidence to continue one's education, and the attempt to measure the development of trust, communication, and concern with the community are measures unique to this particular study.

Another approach for comprehensive evaluation was presented in a study by Joseph and Almen (1970) of a Work Opportunity Center (WOC). The WOC program offered training in a variety of skills (e.g., machine shop, graphic arts, food preparation, etc.) combined with courses to help develop communications, social studies, and mathematics skills. In addition, the student was given an opportunity to earn high school credits at the WOC. Although several standardized and specially developed tests and measures were utilized, most of them were used for diagnosis and prescription rather than evaluation. Educational history, reading level, and family background were some of the types of data gathered on each student. However, some pre and post test data were examined to determine the effectiveness of the program in terms of immediate outcomes. These measures included attitude and self-perception questionnaires developed for the program, an analysis of attendance patterns, and a record of the students' plans for the future. The follow-up studies to assess the more long-term effects included:

- Job success as measured by employer and student evaluation.
- Academic success as measured by the re-entry of the student to a regular high school program.
- The value of the training as measured by a determination of what the graduates were doing six months and one year after leaving WOC. Data were collected on the number and types of jobs secured.

Drane (1971) examined a wide range of variables when assessing project STAY. This project, conducted in St. Louis, Missouri, consisted of a broad approach to the treatment of dropouts. The program consisted of work-study, guidance, continued education, social and personal adjustment, curriculum revision, and after school activity components. The project objectives and measurement techniques are listed below:

Measures	Objectives
1. School records	Reduce the percentage of dropouts
2. DEMOS Dropout Scale	Improve the attitude of dropouts or potential dropouts toward school
3. School records	Reduce absenteeism in school
4. School records	Reduce the number of suspensions in school
5. Questionnaire and social worker visits	Entry of students into post-high school educational programs and/or successful post-high school employment
6. School records	Students continuation in school after pregnancy
7. School records	Reduce rate of failing grades
8. The Iowa Test of Basic Skills for language and arithmetic	Improve achievement in basic skills
9. School records	Have students participate meaningfully in after school activities

Using subjects from two Neighborhood Youth Corps out-of-school projects Freeberg and Riley (1972) attempted to develop and validate an evaluation test battery. The measures used and their characteristics are listed below:

Measure

1. Practical Reasoning Tests - simple Zip coding and map reading tasks.
2. Job Knowledge Test - a multiple choice test on a variety of jobs. It included questions on educational requirements, starting salary, primary tasks performed, hours of work, etc.
3. Job-Seeking Skills Test - a multiple choice test dealing with ways

of looking for jobs.

4. Job-Holding Skills Test - a paper and pencil test in which the respondents must describe the appropriate behavior called for in described situations. The items were designed to elicit responses concerning appropriate dress, punctuality, etc.
5. Vocational Aspiration Minus Plans Test - a measure of the discrepancy between vocational aspirations and vocational plans.
6. Attitudes Scales - To assess attitudes toward authority, self-esteem, motivation for vocational achievement, and willingness to accept deferred gratification.
7. Vocational Interests Scales - To measure the respondents preferences for performing various kinds of tasks such as clerical, service, technical, etc.
8. Standardized Achievement Test (Paragraph meaning sub-test of the Stanford Achievement Test, Form W, Intermediate I; Arithmetic and Vocabulary sub-tests from the Adult Basic Learning Exam Level II, Form A; a Figure Analogy Test; A Maze Following Test; and a Rote Memory Test.)

The measures described above were compared with criterion measures in order to determine their predictive validity. The cirterion measures were divided into two parts- (1) program completion criteria and (2) post-program criteria. The initial criteria for program completion consisted of adjustment to the training program, social adjustment (including adjustment to the community and the family, police contact, peer ratings), and job aspiration level. The last criterion considered salary expectations, the level of the current job and the level of long-range plans. The post program criteria consisted of general job success and adjustment as determined from employer proficiency ratings, financial savings, ability to retain employment, length of stay on job, family adjustment, and police contact. In addition, some assessment of striving for vocational success was made by examining job promotion, amount of first raise, problems with people in the community, the quality of the next job desired, and the ease with which he was contacted by the interviewer. And finally the subject's "stability-mobility" was examined by determining the number of places he interviewed for a job, the number of jobs

held, the visits to state employment service, and the number of places lived. The preliminary data were collected and factor analyzed. There were promising results on some measures. However, the small sample prohibited drawing any significant conclusions from the data.

In evaluating the dropout prevention program designed to work with ninth grade students in Cleveland public schools, Kilbane and Fleming (1972) conducted an assessment of the of the degree to which the program was able to achieve its objectives. The immediate program outcomes and measures are listed below:

Immediate Program Outcomes	Measure
1. A majority (75%) of the students would attain a "normal" rate of progress in reading while in the project	Comprehensive test of basic skills in vocabulary and reading comprehension
2. A majority (75%) of students would attain a "normal" rate of progress in mathematics while in the project	Comprehensive tests of basic skills computation test
3. Students would attain better school marks while in the program than they did prior to participation	Grade point average
4. Students would evidence a higher attendance rate while in the program than they did before entering the program	School records
5. Program students would evidence a lower dropout rate than other ninth grade students in the project schools	School records
6. Students would improve the quality of their written classwork and homework assignments	Teachers' reports
7. Students attitude toward self, school, and teachers would improve	Attitude scales

Project Outreach (1973) was a broad school-wide program to reduce dropouts. The immediate program outcomes and evaluation instruments used are listed below:

Immediate Program Outcomes	Measure
1. Enrollment in school would be maintained during the first project year by 90 of the 100 students identified as potential dropouts.	Official school records
2. Percentage of D and F grades of the targeted students would be decreased by 10%.	Official school records
3. Absenteeism of the targeted students would be decreased by 7%.	Official school records
4. Reading comprehension and vocabulary would be increased by five (5) percentile points for targeted students.	Iowa Test of Educational Development and the Iowa Test of Basic Skills
5. The attitude of targeted students toward adult authority and toward their environment would be improved.	Attitude Survey (Evaluator Design--Semantic Differential Type)

Hornbostel, et. al. (no date) conducted a complex and comprehensive evaluation on a program to assess the effects of various kinds of training programs on the rehabilitation of the dropout. Four groups of dropouts were followed over a two-year period. There were three experimental groups: one received vocational and academic training, the second received vocational training alone, and the third was given academic training alone. The fourth was a control group and received no training. In addition, a fifth (second control) group was composed of students who had started but did not complete any of the experimental programs. Because of attrition, the two control groups were later combined into one. There were 118 subjects in the Academic Vocational Group, 115 in the Vocational Group, 59 in the Academic Group,

and 46 in Control Group I. Unfortunately, as time went on and attrition took place, all of the groups shrank in size. The N's for the various groups at the three testing points are shown in Table I. During the follow-up the N's ranged from 4 to 60.

The research design consisted of:

- PRETEST--the assessment of differences among the group prior to treatment.
- BEHAVIOR DURING TRAINING--this included measurements of the following:
  1. Achievement
  2. Social adjustment
  3. Interests
  4. Anxiety
  5. Social class orientation
  6. Rural-Urban orientation
- POST TEST--the assessment of differences among the groups after training.
- BEHAVIOR AFTER TRAINING--the assessment of clients using the following criteria:
  1. Social adjustment
  2. Interests
  3. Anxiety
  4. Aspirations
  5. Job placement history
  6. Job success
  7. Job satisfaction
- TWO-YEAR FOLLOW UP--an assessment of the differences among the groups two-years after the training program.

The progress and immediate outcomes of primary concern were:

1. Academic achievement
2. Changes in social psychological characteristics
3. Achievement in vocational training

In addition, the long-range assessment criteria were related to the graduate's:

1. Ability to obtain and hold a job
2. Ability to perform the job satisfactorily
3. Continuation of education and/or training

TABLE I

Number of Subjects Tested at Each Test Period

(Hornbostel, et. al.)

N's for the Study

Group	Pretest	Posttest I	Posttest II
Academic-Vocational	118	56	18
Vocational	115	42	13
Academic	59	19	15
Control I	46	15	12
Control II	-	13	14
TOTAL	338	145	72

4. General satisfaction with life

The evaluation instruments and administration schedule are shown in Table II.

The General Aptitude Test Battery yields 9 aptitude test scores and is widely used by the United States Employment Service. The Sequential Test of Educational Progress measures academic achievement in reading, writing, social studies, mathematics and science. The California Test of Personality was used to measure "life adjustment" and is a composite of the "personal adjustment" and "social adjustment" scales. The Kuder Preference Record, Personal Form A, was used to measure interest in social situations related to work. The IPAT Anxiety Scale was used to measure clinical anxiety in an objective and standardized manner. The Social Class Inventory, developed by S. Sutker, is an experimental instrument and is intended to reflect either middle or lower class orientation. The Rural Orientation Inventory is also an experimental instrument authored by S. Sutker.

Other key data collection instruments consisted of the Initial Data Sheet which was developed to obtain basic personal background data, the School Dropout Interview Schedule used to acquire demographic and background information about each subject, and the Youth Opportunity Survey which was administered at six month intervals after training. In addition, the Brayfield-Rothe Job Satisfaction questionnaire and the Goetzel Job Performance Scale were used to measure a subject's vocational success as viewed by both themselves and their employers.

Although the small number of subjects tends to lay open to question some of the findings of this evaluation study, it was undoubtedly one of the most ambitious and well planned of any of the evaluation programs located in the literature. It included various assessment techniques prior to training, during the course of training, and at varying periods following training. In addition the measures covered aptitude, achievement, personality, and values, as well as specific skills required for particular trades and jobs. These were supplemented by structured interviews and ratings scales as described.

TABLE II

Instrument Administration Schedule

(Hornbostel, et. al.)

Instrument	Pretest	During Treatment	Posttest (12 mos.)	Posttest (24 mos.)
General Aptitude Test Battery	X			
Sequential Test of Educational Progress	X		X	X
California Test of Personality	X		X	X
IPAT Anxiety Scale Questionnaire	X		X	X
Social Class Value Orientation Inventory (S. Sutker)	X		X	X
Kuder Preference Record & Personal	X		X	X
Initial Data Sheet Record	X		X	X
School Dropout Research Interview Schedule		X		
Every Pupil Scholarship Test in Typewriting	X		X	
Hiett Simplified Shorthand Test	X		X	
Mullenbruch Office Skills Achievement Test	X		X	
Office Machine Mechanic Test			X	
Purdue Trade Information Test in Welding			X	
Purdue Test for Machinists and Machine Ooperators			X	
Cosmetology Test				X
Mullenbruch Garage Mechanic Test				X
Teacher's Rating of Vocational Skill Performance			X	X
Youth Opportunity/Follow-Up Survey*				
Goetzel Job Performance Scale*				
Brayfield-Rothe Job Satisfaction Blank*				

*Utilized in follow-up study

B. RESULTS FROM PREVIOUSLY USED MEASURES

The data in this section are presented in the general order of the model described in Section III. They range from measures of progress and immediate outcomes to post program objectives and long-range goals. It should be noted that quantitative data were not available for all of the measures previously mentioned.

1. Measures of Progress and Immediate Outcomes

a. Continuing Education-Retention in School

Chance and Sarthory found that out of 64 enrollees in their program, 56 remained and 2 joined another program (Job Corps). "This is a retention rate of 91% which equals the retention rate for the junior and senior high schools at large despite the fact that project enrollees were selected specifically because they were considered most likely to drop-out..." (Chance and Sarthory 1972, p. 213).

Sarthory (1971) reported that of the 71 potential dropouts in a work experience program 26.6% did not complete the year. This compares with a national dropout rate of over 30% and over 40% in some urban school districts.

In Project STAY conducted in the St. Louis Public Schools the Dropout Rate was reduced from 17.86% to 8.97% at one high school and from 1.46% to 0% at a second over a 2 year period. In addition, 96.1% of the known pregnant girls (N=100) at one of the high schools continued in school. Suspensions at the 2 high schools, which had a combined enrollment of over 3500, were reduced from 3.36% to .38% over the same two year period (Drane, 1971).

Kilbane and Fleming (1972) reported a dropout rate of 4.5% for their program as compared with 4.7% for all other 9th grade students in the five project schools.

Project Outreach (1973) reported that of the to al students enrolled in an Enriched Education program (N=215) only 5% dropped out as compared with a 7% overall dropout rate for the school.

b. Continuing Education - Attendance

Davis (1972) reported that absences declined (N=15) from 306 to 50 from the first to second semester. Chance and Sarthory (1972) found a decrease in average absences from 46.6 to 40.9 days after 1 year of the program (N=56).

Hamburger found improved attendance for a group of Coop students (N=108) when compared with a control group (N=91). Average absences fell from 8.09 to 6.09 for the Coop students and went from 7.10 to 7.31 for the controls. Royston (1970) reported a 39% increase in attendance after one year of a special program to keep potential dropouts in school.

Attendance increased from 85.8% to 88.4% for the boys and from 84.4% to 86.3% for girls in the Production Workshop Program for 9th grades. The average for attendance for all 9th grades in the regional school program was 85.6% (Kilbane and Fleming, 1972), while in Project Outreach (1973) absenteeism dropped from 11% to 9%.

Drane (1971) found no significant improvement in attendance for participants in the STAY (dropout prevention) program. However, at one school the percent of students absent 30 days or more dropped from 10.4% to 9.7%.

c. Intellectual Achievement

In a study of students in a cooperative work-study program reported by Hamburger in 1965, it was found that Coop students (N=108) had a mean increase of 5.50 points in IQ vs. 1.55 points for the controls (N=91). There was no significant difference between the groups prior to the program (Coop $\bar{X}$ = 84.0; control $\bar{X}$ = 85.17). The IQ test used was the Otis, Gamma, Form EM.

In a program designed to give a group of dropouts the academic preparation necessary to qualify for a diploma, Kaufman, et al. (1968a) reported a statistically significant increase in average IQ for the experimental group (N=60). Verbal IQ was raised from 92.1 to 94.0 (Otis Quick Scoring Mental Ability Test, Gamma Test Forms EM and FM).

d. Basic Academic Skills - Achievement Tests

Hamburger (1965) found no significant increase in the Stanford Achievement test scores for Word Meaning and Paragraph Meaning in a sample of

108 Coop students vs. 91 control students. He measured the students' pre and post program participation. He also found no significant increase in arithmetic ability as measured by the Stanford Achievement Tests as a result of participation in the Cooperative Work Study Programs. The Coop students were compared with a group of controls (N=91) who also showed no significant gain.

The Public Service Career Program (PSCP) of the City University of New York (1967) reported across-the-board increases in the California Achievement Test (CAT). The PSCP program is aimed at assisting past high school age individuals to prepare for the high school equivalency examination. The CAT grade equivalence rose from 7.8 to 8.0 for reading, 7.2 to 8.4 for arithmetic, and 7.6 to 8.9 Language in only 4 months. A grade of 9.5 is considered the "equivalency" level and appeared to be in reach of the 9 to 12 month program. The Gates Reading Survey was used as a screening device but one recommendation of the progress report was to switch to the Adult Basic Learning Examination (ABLE). It was stated that the ABLE is untimed, is not as culturally biased and appears to be more precise.

Kaufman, et al. (1968a), reporting on an academically oriented drop-out prevention program, found a gain in reading and arithmetic achievement measures as measured by the Stanford Achievement Tests. The Advanced Paragraph Meaning and Advanced Arithmetic Tests, Forms W and X, were used.

Drane (1971) reported no significant improvement in the Iowa Test of Basic Skills (language, arithmetic, vocabulary and reading) for project STAY participants.

In the Cleveland Public Schools Production Workshop Program for potential dropouts in the 9th grade, Kilbane and Fleming (1972) reported that 71% of the participants (N=23) demonstrated "normal" progress in Vocabulary and 56% "normal" progress in Comprehension. This despite the fact that marked deficits in these two areas were measured in pre-test scores (grade 4.4 in Vocabulary and 3.8 in Comprehension). These changes and the Computation test changes are shown below.

	Boys			Girls		
	Vocab.	Comp.	Computation	Vocab.	Comp.	Computation
Pre-Test	4.0	3.5	5.3	4.7	4.0	5.9
Post-Test	5.7	5.0	6.0	5.8	4.9	6.9

Using the Stanford Achievement Tests, Erikson and Hamler (1972), reported increases in both reading and mathematics for students in a "Street Academy." The increase was measured for 20 students over a seven month period. English levels went from 6 yrs. 4 mos. to 7 yrs. 1 mos. while Mathematics increased from 5 year 8 mos. to 6 year 2 mos.

Using the Iowa tests of Educational Development and the Iowa Tests of Basic Skills, Project Outreach (1973) reported that the mean percentile for Reading Comprehension went from 10 to 23 (N=127). These were aggregate scores for program participants grades 7 through 12. No increase was noted, however, for Mathematic skills.

In a program for school alienated youth in Norwalk, Conn. using 13 test variables, only the vocabulary section of the California Reading Test showed a significant increase over the controls. The experimental group improved from 7.4 to 8.3 in grade placement vs. 8.7 to 8.8 for the controls. All were in the 9th grade. Of particular interest, it was additionally found that there were significant correlations between employer occupational ratings and post-test scores in Reading and Speech. It would appear that these verbal skills are related to job success. (A pilot project, 1967).

e. Basic Academic Skills - Grades

Hamburger (1965) found a greater (statistically significant) improvement in class averages for a group of Coop students when compared with a control group. The Coop group did start from a lower baseline however (69.97 to 73.53 for the Coop and 70.69 to 71.95 for the Controls).

Drane (1971) reported a reduction in failing grades (from 13.1% to 11.6%) for project STAY participants while Davis (1972) found that average grades went up for his sample (N=15) from "D" average to "C+" from first semester to second. Project Outreach (1973) reported a 53% decrease in "D" and "F" grades over a one-year period for a group of students (N=25) participating in a dropout prevention program.

Kilbane and Fleming (1972) reported an increase in grade point average from .89 to 1.91 for boys and 1.57 to 2.49 for girls in a dropout prevention program for 9th graders. In a statewide evaluation of a Cooperative Vocational Educational Program in the State of Illinois it was found that the Coop students (N=235) had an average gain of .502 in grade point average, vs. a drop of .066 for the 220 controls (School, community and youth, 1972).

Sarthory (1971) was the only one to report a decrease in grade point averages. His data were from 41 students in a summer work experience program. The greatest decrease was in 14 senior girls which went from 1.75 in the fall of 1969 to 1.13 one year later.

f. Measures of Attitudes and Alienation

Leubling and Trobe (1965) found that as a group program participants tested with the TAT (N=47) tended to:

1. Have difficulty in coping with reality factors including low frustration tolerance, inability to postpone immediate gratification of needs, pessimism, and conflict between environmental demands and personal desires.
2. Inadequate self-awareness and confusion concerning social role requirements.
3. Impaired interpersonal relationships particularly with adults and authority figures.
4. Immaturity and dependency.
5. Ambivalent and negative attitudes toward society.

It was concluded, however, that these characteristics were reflections of experience and values of the sub-culture and were not pathological.

Hamburger (1965) found "... that the disadvantaged youth as a group started with a modest level of educational and vocational aspiration which did not change significantly during the study" (Hamburger, 1965, p. 13). These data were from the Rotter Incomplete Sentence Blank. However, Drane (1971), working in two high schools in St. Louis, showed that a dropout prevention program (Project STAY) could be effective in

improving student aspirations as indicated by the "Student Progress Chart" and the number of participants taking National College Entrance Tests.

With regard to self-concepts Kaufman, et al. (1968a) reported an increase in feelings of self-esteem during the course of an academically oriented dropout prevention program. Attitudes were measured by the Gough Adjective Check List, the Coopersmith Self-Esteem Inventory and the Edwards Personal Preference schedule. Using an Adjective Check List self-concept instrument, Joseph and Almen (1970) investigated changes in self-attitudes before and after participation in a Work Opportunity Center Program. In one sample of 54 students (average age 18 years, 1 month) who had been enrolled over a period of 16 weeks, approximately 60% of the group made improvements in self-perception scores. Another sample of 110 students (average age 17.4 years) who were ready to leave the program (to return to school, start jobs, or graduate) showed a significant increase in positive items. The males showed a greater increase in positive items than the females.

Attitudes toward school and authority were also measured using a variety of instruments. Using an attitude scale and observations by consultants and evaluation specialists, Erickson and Hamler (1972) reported positive attitudes and lack of alienation toward school among 42 participants in a dropout prevention program. These students were all in the program because they had initially dropped out of school.

Drane (1971) found that after participation in project STAY, 84.4% of the participants at one high school and 81.5% at the other had a positive attitude toward school as measured by the Demos Dropout Scale. There were no baseline scores for comparison. In addition, advisors perceived 80% of the students as having a positive attitude toward school. However, the goal of having at least 50% of the unemployed students participate in after school activities as an indication of positive attitudes was not realized. Using an evaluator designed Semantic Differential Attitude Survey, Project Outreach (1973) reported improvement in student's attitudes toward adult authority and to their environment. Davis (1972) found that after 1 year of participation in a dropout prevention program (N=223) 93% felt their attitude toward school improved "at least a little." Most of the participants felt that having "work experience" gave them an

increased feeling of responsibility at home (99%), at school (97%), and at work (100%).

Pre and post attitude measures on program participants (243 ninth graders) were taken in three areas -- "self as student", "social aspects of school" and "perception of teachers." Only negligible changes were reported in the boys' responses and these reflected a decline in positive feelings. The girls responses, however, reflected an increase in positive feelings (Kilbane and Fleming, 1972).

Sarthory (1971) reported on a summer employment program involving work exploration, remedial education, and counseling for 71 students. He found that attitude toward teachers improved and school appeared more desirable and interesting but less useful at the end of the program. The latter is perhaps due to the discrepancy between what went on in school and the real expectancies of employers.

g. Acquisition of Social Skills

Although the social requirements of work are a major contribution to job success or failure (Clark, et al., 1969), virtually nothing has been done to evaluate this area. The one study located in the literature found that, for those students who were unsuccessful in retaining employment once placed in a job, reports from their employers indicated that the cause tended to be lack of desired work attitudes rather than lack of technical skills (Joseph and Almen, 1970).

2. Post Program Objectives and Long-Range Goals

a. Continued Education and Employment

Erickson and Hamler (1972) reporting on the Benjamin Franklin Urban League Street Academy in New York City found that 84% of the students from the 1970-71 class upon whom data were available (N=58) were working or in school. Of the 1971-72 class, 94% of the 110 students were still in school or working. They identified continuing one's education as a key evaluation tool. It was hypothesized that continuing participation in an educational program reflected a self-confidence to achieve in a school setting, and as such would be a behavioral index of such self-confidence. It was found that 72% of the 1971-72 sample (N=136) did in fact remain in school.

Joseph and Almen (1970) found that of those who attended the Work Opportunity Center, 53% had earned an average of 4.9 academic credits each while a matched group of non-attendants made no academic progress. Of 260 respondents to follow-up questionnaires in project STAY (St. Louis) 36.5% were employed and 47.3% were in post-high school educational programs. Of those surveyed from the Work Study portion of the program (N=35), 28.6% were employed and 51.4% were in educational programs. Of those employed, however, only 30% were in occupations related to the Work Study Training (Drane, 1971).

Weber (1972), in a follow-up study of 41 student graduates from a "continuation high school" in the San Francisco Bay area, found that 60% of the boys were employed full-time, one-third of the boys were enrolled in trade schools or apprentice-programs, and only two boys were unemployed. "Two-thirds of the girls were married"; they constituted the group "not employed full-time." (Weber, 1972, p. 573). In 2 years of a project on upgrading the employability of dropouts (N=3285, ages 16-20), McCarthy (1970) reported that 751 (23%) achieved full time employment and 714 (22%) were enrolled in a regular full-time school or a training program.

Over a 45 month period with their Work Opportunity Center in Minneapolis, Joseph and Almen (1970) found that 1250 of the 2761 enrollees (45%) were placed in jobs. In a follow-up study (elapsed time unspecified) of 236 students placed in jobs "about 60% appear to have made an adequate adjustment at their work stations... Another 30% had moved on to another job or had left the city. The remaining were seeking new jobs or were unemployed." (Joseph and Almen, 1970, p. 80). They also found in their follow-up study that those who attended the Work Opportunity Center earned an average of $.15 per hour more than a matched group of non-attendees (1.75/hr vs. 1.60/hr).

Hornbostel, et al. (no date) reported on a study comparing three groups of dropouts with a control group. The three dropout groups each had different training: Academic only, Academic and Vocational and only Vocational. They reported no significant differences among any of the groups with regard to job satisfaction, employer ratings, proportion continuing in their education or training, or in the extent of their general satisfaction with life. Over the two year follow-up they did find

that the Academic-Vocational and Vocational group females were significantly higher than the controls with regard to length of tenure of employment and average earnings.

b. Citizenship and Community Concern

When compared with the control group Walther (1967) reported that Black female participants in four urban Neighborhood Youth Corps Program were more frequently employed and self-supporting. There was a significant drop in police contacts at two of the sites. The males showed significant improvement associated with continued education, both vocational and academic. One possible explanation for the favorable findings with respect to Black females is that they "... are easier to work with than are males, they have more interest in improving themselves, and they have a more difficult time obtaining job training or job placement without assistance" (Walther, 1967, p. 4). Sarthory (1971) also found that the summer work experience program successfully reduced the crime rate and police encounters among participants.

Erickson and Hamler (1972) examined "Trust, Communication and Concern with Community" as an index of self-development. While there were no quantitative data available it was felt that the drop-out prevention program did "... reinforce the development of consciousness and concern for the community of the students." (p. 14).

c. Other Long-Range Effects

Corazzini, et al. (1966) examined the benefits and costs of vocational education in Worcester, Massachusetts. He concluded that the vocational school was not impressive as a dropout prevention mechanism. The vocational school graduates were not favored over regular high school graduates by local firms and although the vocational school aided graduates in job placement, their career paths inside the firm were the same as those of regular high school graduates. Further, the vocational schools did not aid intra-regional movement (mobility) of workers. The only positive outcome appeared to be "inter-generational" movement of workers (upward mobility). However, this will depend on whether the students actually follow the career paths they have chosen them.

V. SUMMARY AND CONCLUSIONS

An extensive review of the literature has revealed little in the way of evaluation approaches, techniques, or measures that are not currently incorporated or accounted for in the existing evaluation system designed by Gibboney and Associates. The objective of the current modest effort was to attempt to identify criterion-referenced measures appropriate for the assessment of both the immediate and long-range objectives of the OIC/CIP program. While we were able to identify several similar programs in the literature, there have been few attempts to systematically evaluate the outcomes of these programs to the extent currently being attempted by Gibboney and Associates. For the most part the descriptions of the evaluation plans, if described at all, tended to be superficial. In some cases where there were attempts at devising a more comprehensive in-depth and long-term evaluation, the data presented in the report itself were sketchy. It was difficult to tell from the reports what was done in the way of evaluation, let alone what the outcomes were. In summary, the literature in this area is quite sparse with little quantitative data.

With regard to the measures themselves, since there was little reported on criterion-referenced measures, this report was broadened to include all of the measures identified in the literature that might prove useful. These ran the gamut from standardized tests to program specific tools and from objective IQ tests to projective personality measures. There were basically two general types of measures used. There were those measures which reported actually observable behavior and those which were tests or indirect measures. The directly observable behavior measures included items like attendance, holding a job, not becoming involved with the police, etc. It was of course assumed that the indirect measures were related to and precursors of behavioral change. However, there was no real evidence presented that this was indeed the case.

Unfortunately few of the researchers attempted to use measures employed by others, so it is difficult to draw any conclusions regarding their general utility or applicability. (If a test is used in only one program and no change is recorded, does this mean the measure is insensitive or

the program is not working?) The one type of test used most frequently was the standardized achievement test - particularly for reading and math. The Stanford, Iowa and California Tests were the most common. In addition another commonly used set of measures were those used to periodically assess self-concept. These measures included adjective checklists, sentence completion forms, and student self-descriptor forms. In addition, structured interview forms were used by counselors for these purposes as well, either as a substitute or in addition to the self-evaluation.

Other scales and checklists were used to assess attitudes toward authority, work, and school. It was thought that changes on these attitudes might be precursors of program success since they appear to be highly related to continuing in school and to successful employment. As the OIC/CIP staff pointed out during one of the site visits, a major problem appears to be getting the student "turned around". Several of the programs included personality tests assuming that there would be some positive program effect on such factors as emotional maturity, emotional stability, and hostility, etc. They also tried to obtain from these measures an indication of delayed gratification, like willingness to postpone monetary and other kinds of rewards while completing one's education and training. The attitude or personality measures also reflected the degree of feeling of control over one's future and destiny (the degree to which one feels he has something to say about what happens to him as opposed to the idea that everything that happens is predetermined by where he is and who he is). Also, personality measures have been used as indicators of motivation, ambition, desire to achieve, aspiration level, and desire for responsibility. However, there appeared to be nothing reported in the literature which would aid in determining the maturity with which career decisions are made.

Behavioral measures included the obvious things like the percentage of students that stay in school, number of encounters with police, grades, absenteeism, etc. The criterion-referenced measures of program success were tied directly to the major objectives of the programs. These included: 1) Keeping the student in school, 2) Having the student achieve at least GED or high school diploma, and 3) Having the student enter additional training or an entry level position for which he has been trained in the

program, or 4) Having the student enter college. As a long-range follow-up measure there were indices of job holding skills and "job success" as measured primarily by interview forms and rating scales.

There were few measures, however, directed at assessing the basic skills required for job application, such as what to do on an interview, how to write a letter of application, how to talk on the telephone and make an appointment, etc. This despite obvious importance of these skills to obtaining and holding a job.

The fact that so many different measures were used would seem to imply that no one was really satisfied with any given set of evaluation techniques. It would appear that the data gathered with these measures were not providing adequate information for the assessment of program effectiveness. The measures found in the literature were for the most part related to the specific short-term expectations of each particular program. They were based on immediate program outcomes and were not empirically tied to the ultimate long-range goals. Thus, the data obtained consisted of many individual scores and indices which presented only an incomplete and fragmented picture of program progress and success. It is apparent that what may be desirable is an entirely new approach to program evaluation based on the answer to the basic question "What have we really done for this individual?" Is the individual really better off for having been a product of the program? Such an assessment is neither summative nor formative but rather is oriented toward impact and ultimate payoff. It is not asking the question of whether the objectives of the program have been met but rather if the individual participant has improved his potential for a better life, i.e. were they the correct objectives?

This approach to evaluation has at least one serious drawback. It is often difficult to relate any long-term change in the individual's potential with the direct outcome of the program itself. However, the advantage to such an evaluation approach is that the evaluation becomes much more meaningful in terms of real and tangible benefits to the individual. There is virtually nothing in the literature having to do with this kind of approach to assessment. There are no precedents, no measures, and virtually no conceptual framework within to operate. So while the question appears simple--how have we improved the life of this human

being--the approach to answering such question is highly complex indeed.

Such an assessment approach as we are describing is centered around the individual, and is criterion-referenced and socially oriented. In addition, the measures should ideally be program independent, i.e., independent of the stated program objectives. For example, the objectives stated in the program are having this individual obtain a high school diploma or a GED, obtain hands-on experience in several occupational areas, obtain an entry level job or training position, etc. The question still remains-that if the students obtain any or all of the stated objectives of the program, are they better off than they would have been if they had not attended the program; and are they better off than other people with similar backgrounds who have not attended the program? The impact assessment is based on a critical analysis of the assumptions underlying the program and the validity of its objectives. It attempts to determine if the accomplishment of the objectives contributes to the improved welfare of the individual.

This aspect of the evaluation then takes the form of examining the program independently and objectively from the outside. It starts with the end goal and works back toward the program itself. The first step will need to be the establishment of some criteria by which the success of the program can be assessed over a long-range period. These same long-range measures can also be applied at intermediate stages to determine what progress is being made. In addition, intermediate measures need to be identified which are predictive of the ultimate criterion. These may include some of the more traditional tests and measures identified and some new measures developed specifically for this purpose. For example, one possible measure for assessing the ultimate impact of the program may be the general concept of "investment". In an AIR impact evaluation conducted overseas (Krug and Jung, 1974), it was proposed that one measure of the individual's "engagement" with society (as opposed to alienation from) is the amount of time, money and effort he is willing to give to his community. This may be reflected in volunteer work, charity contributions, political activity, etc. Using the same measure one can assess an individual's willingness to invest time, money and/or effort in his own continued self development, e.g., continued education, travel, participation, in cultural events, etc.

In addition, interim progress measures need to be developed which are related to these impact measures. For example one possible progress measure related to ultimate program success may be a relationship between geographic mobility in job-seeking behavior and maturity and self-concept. It may be that the more mature, self-confident individual will be willing and able to seek jobs or employment outside of his normal sphere of operation. This allows for greater choice and a better probability of finding suitable employment.

In the CIP program being run by the OICs of America, the ultimate goal is both unambiguous and obvious--to improve the social and economic well-being of those individuals participating in the program. Any measures of program success must center around the individual's social and economic development. The interim or progress measures should reflect the growth, or lack of growth, exhibited by the individual from the time he enters the program to some preselected end point. It is suggested that this period of evaluation may well extend for many years in the future. To fulfill this aspect of the evaluation a considerable effort needs to be undertaken to develop and validate both the ultimate and intermediate progress measure. Such an effort would not only benefit the current program, but would contribute unmeasurably to the advancement of the state-of-the-art in evaluation.

APPENDIX A

MEASURES	REFERENCES
A. Intelligence	
Culture Fair Intelligence Test, Scale 2 Forms A & B	(Kaufman et. al., 1968a)
Lorge Thorndike Intelligence Test	(A pilot project-occupational training, 1967; Beecher and DiPasquale, 1962)
Otis Intelligence Test, Gamma Form E	(Hamburger, 1965)
Otis Quick-Scoring Mental Ability, Forms EM & FM	(Leubling and Trobe, 1968a)
Weschler Adult Intelligence Scale	(Leubling and Trobe, 1965)
B. Aptitude	
Armed Forces Qualifying Test	(Savitzky et. al., 1965)
Airman Qualifying Examination	(Savitzky et. al., 1965)
Bennett Mechanical Comprehension Test, Form AA	(Sharar et. al., 1969)
Civil Service Examination	(Exemplary vocational education program-Final Report, 1973b; Savitzky et. al., 1965)
Differential Aptitude Test	(Beecher and DiPasquale, 1962; Hamburger, 1965)
General Aptitude Test Battery	(Dayton, 1973; Exemplary vocational education program-Iterim Report, 1973a; Horbostel et.al., no date; Savitzky et. al., 1965; Thurston and Weber, 1971)
C. General Achievement and Skill Measures	
Agency Intern Evaluation Form	(Off-campus experience...for high schoolers, 1972)
Botel Reading Inventory	(Frost and Pilgrim, 1969)
Career Development Achievement Test	(Randolph and Holmes, eds., 1972)
Crites Vocational Development Inventory	(Randolph and Holmes, eds., 1972)
Cummulative Record	(Erickson and Hamler, 1972; Kaplan 1967)

C. General Achievement and Skill Measures (Continued)

The Diagnostic Reading Test	(Career College, 1969)
Differential Aptitude Test, Battery Form A, Clerical Test	(Hamburger, 1965)
Dolch Basic Sight Word Text	(Frost and Pilgrim, 1969)
Every Pupil Scholarship Test in Typing I & II	(Hornbostel et.al., no date)
Gates Macginitie Reading Test	(Decker and Anderson, 1969; Frost and Pilgrim, 1969; Sharar et. al., 1969)
Gilmore Oral Reading	(Frost and Pilgrim, 1969)
Goertzel Job Success Rating Scale, Form A	(Hornbostel et. al., no date)
Grade Point Average	(Chance and Sarthory, 1972; Davis, 1972; Evaluation of the in-school youth work-training project, 1968; Lowe et. al., 1970; Project Out-Reach, 1973; Royston, 1970; Sarthory, 1971)
Heitt Simplified Shorthand Test	(Hornbostel et. al., no date)
Iowa Achievement Tests	(Project Outreach, 1973)
Job Corps Reading and Math Placement Tests	(Twelfth month evaluation-Project Interchange, 1968)
Job Ratings	(Off-campus experiences...for high schoolers, 1972)
Kottmyer Diagnostic Test of Word Perception Skills	(Frost and Pilgrim, 1969)
McCall-Crabs Standard Test	(Frost and Pilgrim, 1969)
Mellenbruch Office Skills	(Hornbostel et. al., no date)
Metropolitan Achievement Tests	(Bitter, 1966; Sharar et. al., 1969; Twelfth month evaluation-Project Interchange, 1968)
Morrison-McCall Spelling Scale	(Frost and Pilgrim, 1969)
New York University Speaking Test	(A pilot project- occupational training, 1967)

C. General Achievement and Skill Measures (Continued)

O'Connor Finger and Tweezer Dexterity Test	(Leubling and Trobe, 1965)
Office Machine Mechanical Test	(Hornbostel et. al., no date)
Peabody Picture Vocabulary Test	(Joseph and Almen, 1970)
Purdue Tests-Welding, Sheetmetal Work, Machinists and Machine Operation	(Hornbostel et. al., no date)
Questionnaire Survey	(Weber, E., 1972)
Science Research Associates Phonic Survey	(Frost and Pilgrim, 1969)
Sequential Tests of Educational Progress, Forms 3A & 3B	(Hornbostel et. al., no date)
Stanford Achievement Test	(Beecher and DiPasquale, 1962; Erickson and Hamler, 1972; Frost and Pilgrim, 1969; Hamburger, 1967; Kaplan, 1967; Kaufman et. al., 1968a)
Teacher Progress Reports	(Hornbostel et. al., no date; Kaplan, 1967)
Vocational Training Ratings by Instruction	(Sharar et. al., 1969)

D. Personality

California Psychological Inventory	(Thurston and Weber, 1971)
California Test of Personality, Forms AA & BB	(A pilot project-occupational training, 1967; Hornbostel et. al., 1969)
Case Study	(A pilot project-occupational training, 1967; Savitzky et. al., 1965)
Citzenship Grade Point Average	(Yunker, 1967)
Counselor Evaluation	(Erickson and Hamler, 1972)
Edwards Personal Preference Schedule	(Kaufman, 1968a)
Gough Adjective Checklist	(Kaufman, 1968a)
IPAT Anxiety Scale Questionnaire	(Hornbostel et.al., no date)

D. Personality (Continued)

Potential Dropout Instrument (PDI)	(Randolph and Holmes, eds., 1972)
Rotter Incomplete Sentences Blank, High School Form	Hamburger, 1965)
Self Directed Search Specimen Set	(Korizek, 1972)
Self Esteem Inventory	(Randolph and Holmes, eds., 1972)
Sociometric Form	(Randolph and Holmes, eds., 1972)

E. Attitudes and Interests

Attitude Survey and Observation	(Lowe, 1970; Project Outreach, 1973)
Brayfield-Rothe Job Satisfaction Questionnaire	(Hornbostel et. al., no date)
Business and Industry Job Survey	(East and Dolan, 1968b)
Occupational Adjustment Rating Scale	(A pilot project- occupational training, 1967)
Ohio Vocational Interest Survey	(Dayton, 1973; Exemplary vocational education-Interim, 1973a; Korizek, 1972)
Perceptions of Career Education	(Korizek, 1972)
Questionnaire Survey	(Commuter Job Corps-Final Report, 1970; Davis, 1972; Dayton, 1973; East and Dolan, 1968a; Erickson and Hamler, 1972; Hornbostel et. al., no date; Sarthory, 1971)
Rural-Urban Orientation Inventory	(Hornbostel et. al., no date)
Semantic Differential	(Kaplan, 1967)
Skills Program vs. Diploma Program	(Kaufman et. al., 1968a)
Social Class Value Orientation	(Hornbostel et. al., no date)
Staff Reaction Form	(Kaplan, 1967)
Student Opinionnaire for Career Education perceptions	(Korizek, 1972)
Teacher Observation	(Randolph and Holmes, eds., 1972; Thurs[illegible]nd Weber, 1971)

E. Attitudes and Interests (Continued)

Vocational Sentence Completion Blank	(Cangemi, 1964)
Weingarten Picture Inventory	(Sharar et. al., 1969)
The Work Values Inventory (Super)	(Hamburger, 1975)
California Study Methods Survey	(A Pilot project-occuaptional training, 1967)
Case Study	(Savitzky et. al., 1965)
Demos Dropout Rating Scale	(Randolph and Holmes, eds., 1972)
Internal-External Control Scale	(Career College, 1969)
Interviews	(East and Dolan, 1968a; Erickson and Hamler, 1972; Exemplary vocational education program-Final Report, 1973b; Hamburger, 1965; Hornbostel et. al., no date; Kaplan, 1967; Twelfth month evaluation-Project Interchange, 1968)
Kuder Preference Record, Form A	(Beecher and DiPasuqale, 1962; Exemplary vocational education-Interim Report, 1973a; Hornbostel et. al., no date; Korizek, 1972; Yunker, 1967)
Kuder Preference Record, Form DD	(Korizek, 1972)
Life Planning Questionnaire	(Hamburger, 1965)
Milwaukee Student Survey	(Career Development Program, 1972)
Minnesota Vocational Interests Inventory	(Kaufman et. al., 1968a; Sharar et. al., 1969)

F. Other Measures

Hamburger Revised Occupational Scale for Measuring Socio-Economic Status	(Hamburger, 1965; Leubling and Trobe, 1965)
Keystone Visual Survey	(Frost and Pilgrim, 1969)
Medical Report Tallies	(Kaplan, 1967)
Police Encounters	(Chance and Sarthory, 1972)
Pupil Attendance Record	(A Pilot project-occupational

F. Other Measures (Continued)

	training, 1967; Chance and Sarthory, 1972; Davis, 1972; Hamburger, 1965; Kaplan, 1967; Project Outreach, 1973; Royston, 1970; Yunker, 1967)
Pupil Holding Power Data Instrument	(Cangemi, 1964)
Retention Rate (% staying in school)	(Chance and Sarthory, 1972; Erickson and Hamler, 1972; Project Outreach, 1973)
Reduction of Crime Rate	(Sarthory, 1971)
School Dropout Research Interview Schedule	(Hornbostel et. al., no date)

References

A pilot project to develop a program of occupational training for school alienated youth. Interim Report and Statistical Evaluation, 31 August 1967.

American Institutes for Research. Evaluative Research: Strategies and Methods. Pittsburgh, Pennsylvania, 1970.

Beecher, & DiPasquale. School to Employment Program. Annual Progress Report 1961-1962, 1962.

Bitter, J. A. Training guide for vocational habilitation. 1966.

Cangemi, J. P. A school to Employment Program in Syracuse, New York. 1964.

Career College: A report of the first year (August 1968-June 1969). 1969.

Career Development Program. A report of the pre-program development study concerning the Milwaukee Career Development Program. 1972.

Chance, B. J., & Sarthory, J. A. Dropout prevention: A model for educational change. High School Journal, 1972, 55(5), 208-218.

Clark, J. B., Lichtenstein, S., & Spector, P. Social Requirements of Work. American Institutes for Research. Final Report, AIR-F44-4/69 FR. 1969.

Commuter Job Corps. Final Report, 1970.

Corazzini, A. J., et al. Vocational education, a study of benefits and costs. (A case study of Worchester, Massachusetts). 1966.

Davis, J. A. A state work experience program with multi-agency involvement. 1972.

Dayton, C. W. Coordinated vocational and academic education North Gwinnett High School, Sawanee, Georgia, 1973. Case Studies in Practical Career Guidance, 1973, No. 6.

Decker, G., & Anderson, F. L. Reading, course description. 1969.

Drane, H. T. Dropout prevention program; Project Stay, 1970-71. Final Evaluation Report, 1971.

East, L., & Dolan, M. A. Student performance requirements: Other educational situations. Continuation Education System Development Project. Technical Report 1.3, 1968a.

East, L., & Dolan, M. A. Student performance requirements: Employment situations. Continuation Education System Development Project. Technical Report 1.4, 1968b.

Erickson, E., & Hamler, B. Final report of the evaluation of the 1971-72 Benjamin Franklin Urban League Street Academy. 1972.

Evaluation of the in-school youth work-training project for fourteen and fifteen year old youth, 1968.

Exemplary vocational education program based on environmental studies, K-14. Final FY2 Evaluation Report (Interim Project Report), 1973.

Freeberg, N. E., & Reilly, R. G. Validation of a test battery for youth-work training program enrollees. Research Memorandum, 1972.

Frost, J. L., & Pilgrim, G. H. Reading achievement of Gary Job Corps enrolled in a diagnostically based program, 1967-68. 1969.

Hamburger, M. Report of the evaluation study of the municipal cooperative education program. 1965.

Hornbostel, V. O., et al. The rehabilitation of school dropouts in Oklahoma City: An experimental appraisal. Final Report, no date.

Joseph, M. P., & Almen, R. E. A Work Opportunity Center for Minneapolis, Minnesota. Final Report, 1970.

Kaplan, L. L. Occupational training centers for 16-18 year old youth. A demonstration program for potential or actual dropouts. 1967.

Kaufman, J. J., et al. The school environment and its influence on educational programs for high school dropouts. Interim Report, 1968.

Kilbane, M., & Fleming, M. Production Workshop Project, DPPF; 1971-72 evaluation. 1972.

Korizek, W. A. The next step--a comprehensive program in occupational preparation and placement. Second Year Interim Report, 1972.

Krug, R. E., & Jung, S. M. Systems for Evaluating the Impact of Rural Development Programs. AIR Final Technical Report. AID Contract 493-037R, 1974.

Leubling, H. E., & Trobe, A. S. Vocational Advisory Service Youth and Work Program. 1965.

Lowe, P. K., et al. Challenging inner city potential dropouts. 1970.

Off-campus experiences now an approved course of study for high schoolers. Nation's Schools, 1972, 90(4), 58-60.

Project Outreach for optimum growth of alienated youth who reject school. 1973.

Public Service Career Program (PSCP). Progress Report, October 1967 through April 1968: Education component of the PSCP. 1968.

Randolph, D. L., & Holmes, W. R. (Eds.) A program for early detention and attitudinal reorientation of potential school dropouts. 1972.

Royston, W., Jr. Forsyth County Vocational High an investment in youth. 1970.

Sarthory, J. A. NYC Summer Pilot Project. Preliminary, Interim, and Final Reports, 1971.

Savitzky, C., et al. Developing work-study programs for potential dropouts: A manual. STEP--The School to Employment Program, 1965.

School, community and youth. Statewide evaluation of Part G Program in cooperative vocational education in the State of Illinois/1971-1972. State of Illinois, 1972.

Seeman, M. Alienation and engagement, in Campbell, A., & Converse, P. E. The Human Meaning of Social Change. New York: Russel Sage Foundation, 1972.

Sharar, P. H., et al. The Try Project: A demonstration of a comprehensive educational model for disadvantaged youth. Final Report, 1969.

Thurston, R. L., & Weber, L. C. Development of a plan to improve the personalized instruction program at vocational villages. Final Report, 1971.

Tiedeman, D. V., & Miller-Tiedeman, A. Career Institution When Alienation from Secondary Schools Occurs: A Review of the Literature. American Institutes for Research: Washington, D.C. 1974.

Twelfth month evaluation, Career Planning Center (Project Interchange), September 1967-August 1968. 1968.

Walther, R. H. A summary of retrospective studies of the effectiveness of NYC out-of-school programs in four urban sites. 1967.

Weber, E. J. The dropouts who go to school. 1972.

Yunker, J. A. Pre-high school vocational group guidance for potential dropouts and non-college-bound students, 1967.